

# Detection and Diagnosis of Plant-wide Oscillations From Industrial Data using the Spectral Envelope Method <sup>★</sup>

Hailei Jiang<sup>a</sup>, M.A.A Shoukat Choudhury<sup>b</sup>, Sirish L. Shah<sup>a,★★</sup>

<sup>a</sup>*Dept. of Chemical and Materials Engg., Univ. of Alberta, Edmonton, AB, Canada, T6G 2G6*

<sup>b</sup>*Matrikon Inc., Edmonton, AB, Canada, T5J 3N4*

---

## Abstract

Plant-wide oscillations are common in many processes. Their effects propagate to many units and may impact the overall process performance. It is important to detect and diagnose the cause of such oscillations in order to rectify the situation. This paper proposes a new procedure to detect and diagnose plant-wide oscillations using routine operating data. A technique called spectral envelope is used to detect oscillations. The variables that have common oscillations are identified and categorized accurately by a statistical hypothesis test. A new index called the oscillation contribution index (*OCI*) is proposed to isolate the key variables as the potential root cause candidates of the common oscillation(s). Two industrial case studies are presented to demonstrate the utility and practicality of the proposed procedure.

*Key words:* Oscillation Detection, Diagnosis, Spectral envelope, Principle component, Hypothesis test, Valve stiction, Control loop,

---

## 1 Introduction

Detection and diagnosis of plant-wide disturbances is an important issue in many process industries (Qin 1998, Desborough and Miller 2001). Oscillations are a common type of plant-wide disturbance and the root causes can be

---

<sup>★</sup> A condensed version of this paper has been accepted for presentation at IFAC-ADCHEM 2006, Gramado, Brazil, April 2-6, 2006.

<sup>★★</sup>Corresponding author. Tel.:+1-780-492-5162; fax:+1-780-492-2881. E-mail address: sirish.shah@ualberta.ca

poorly tuned controllers, process or actuator non-linearities, oscillatory disturbance *etc.* The oscillation effects can propagate to many units and thus may impact the overall process performance. The presence of oscillations in a plant increases the variability of the process variables and thus may cause poor control performance, inferior quality products and larger rejection rates. Increasing emphasis on plant safety and plant profitability strongly motivates the search for techniques to detect and diagnose plant-wide oscillations.

Thornhill and Hägglund (1997) used zero-crossings of the control error signal to calculate integral absolute error (IAE) in order to detect oscillation in a control loop. This method has poor performance in the cases of noisy error signals. Miao and Seborg (1999) suggested a method based on the auto-correlation function to detect excessively oscillatory feedback loop. The auto-covariance function (ACF) of a signal was utilized in Thornhill *et al.* (2003a) to detect oscillation(s) present in a signal. This method needs a minimum of five cycles in the auto-covariance function to detect oscillation, which is often hard to obtain, particularly in the case of a long oscillations (e.g., an oscillation with a period of 400 samples). Although the data set can be down-sampled in such cases, downsampling may introduce aliasing in the data. More recently Thornhill *et al.* (2002) proposed spectral principal component analysis (SPCA) to detect oscillations and categorize the variables having similar oscillations. This method does not provide any diagnosis of the root cause of the oscillation which is generally the main objective of the exercise. In this paper, a new procedure based on the spectral envelope method for detection and diagnosis of plant-wide oscillations is proposed.

The spectral envelope method is a frequency domain technique that was first introduced by Stoffer *et al.* (1993) to explore the periodic nature of categorical time series. The idea is to assign numerical values to each of the categories followed by a spectral analysis of the resulting discrete-valued time series. In 1997, McDougall *et al.* (1997) extended the concept of spectral envelope to real-valued series. In exploring the periodic nature of a real-valued series, one can not only do spectral analysis of the original series, but also of transformed series. The key idea in McDougall *et al.* (1997) is to select optimal transformations of a real-valued series that emphasize any periodic nature in the frequency domain. This paper extends the idea of spectral envelope to towards the analysis of process data. The main contributions of this paper are: (1) The concept of spectral envelope is simplified and formulated for use with process data. (2) A statistical hypothesis test is formulated to categorize the variables that have common oscillations. (3) Based on the spectral envelope method, a new oscillation contribution index (*OCI*) is proposed to diagnose the root causes of the oscillations. (4) A novel procedure to detect and diagnose plant-wide oscillations is proposed and illustrated by application on two industrial case studies. (5) The major advantage of the method proposed here is that it can be automated easily.

## 2 Oscillation Detection using the Spectral Envelope Method

In this section, the concept of spectral envelope is introduced. A simulation example is presented to demonstrate its ability to detect multiple oscillations. The performance comparison with the SPCA method is also included.

Throughout this paper, we use (1) bold capital letters to represent matrices; (2) bold lower case letter to represent vectors; and (3) regular letters to represent scalars.

### 2.1 Definition of the Spectral Envelope

Here we provide a simple interpretation of the concept of spectral envelope.

Let

$$\mathbf{x}(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_m(t) \end{bmatrix} \quad t = 0, \pm 1, \pm 2, \dots,$$

be a multivariate, vector-valued time series on  $\mathfrak{R}^m$ . Denote matrix  $\mathbf{X}$  as

$$\mathbf{X} = [ \cdot \cdot \cdot \mathbf{x}(t-1) \mathbf{x}(t) \mathbf{x}(t+1) \cdot \cdot \cdot ]$$

Further, we denote the covariance matrix of  $\mathbf{X}$  as  $\mathbf{V}_{\mathbf{X}}$  and the power spectral density (PSD) matrix of  $\mathbf{X}$  as  $\mathbf{P}_{\mathbf{X}}(\omega)$ . Here,  $\omega$  represents frequency and is measured in cycles per unit time, for  $-1/2 < \omega \leq 1/2$ . The definition of PSD matrix can be found in Jenkins and Watts (1968).

Let  $g(t, \boldsymbol{\beta}) = \boldsymbol{\beta}^* \mathbf{x}(t)$  be a scaled series from  $\mathfrak{R}^m$  to  $\mathfrak{R}$ , where  $\boldsymbol{\beta}$  is a  $m \times 1$  column vector which may be real or complex. The  $*$  represents conjugate transpose of the variable. Actually,  $g(t, \boldsymbol{\beta})$  is a linear combination of the rows of  $\mathbf{x}(t)$ . The variance of  $g(t, \boldsymbol{\beta})$  can be expressed as  $V_g(\boldsymbol{\beta}) = \boldsymbol{\beta}^* \mathbf{V}_{\mathbf{X}} \boldsymbol{\beta}$ , and the power spectral density of  $g(t, \boldsymbol{\beta})$  can be expressed as  $P_g(\omega, \boldsymbol{\beta}) = \boldsymbol{\beta}^* \mathbf{P}_{\mathbf{X}}(\omega) \boldsymbol{\beta}$ .

The spectral envelope of  $\mathbf{X}$  is defined as:

$$\lambda(\omega) \triangleq \sup_{\boldsymbol{\beta} \neq 0} \left\{ \frac{P_g(\omega, \boldsymbol{\beta})}{V_g(\boldsymbol{\beta})} \right\} = \sup_{\boldsymbol{\beta} \neq 0} \left\{ \frac{\boldsymbol{\beta}^* \mathbf{P}_{\mathbf{X}}(\omega) \boldsymbol{\beta}}{\boldsymbol{\beta}^* \mathbf{V}_{\mathbf{X}} \boldsymbol{\beta}} \right\} \quad (1)$$

where  $-1/2 < \omega \leq 1/2$ . It is worthwhile to note that  $P_g(\omega, \boldsymbol{\beta}) = P_g(-\omega, \boldsymbol{\beta})$ , and the relationship between  $P_g(\omega, \boldsymbol{\beta})$  and  $V_g(\boldsymbol{\beta})$  is:

$$V_g(\boldsymbol{\beta}) = \int_{-1/2}^{1/2} P_g(\omega, \boldsymbol{\beta}) d\omega = 2 \int_0^{1/2} P_g(\omega, \boldsymbol{\beta}) d\omega \quad (2)$$

The quantity  $\lambda(\omega)$  represents the largest portion of power (or variance) that can be obtained at the frequency  $\omega$  for any scaled series. The scaling vector that results in the value  $\lambda(\omega)$  is called the optimal scaling vector at frequency  $\omega$ , which is denoted as  $\boldsymbol{\beta}(\omega)$ . Accordingly, the elements of the optimal scaling vector are called the optimal scalings. The optimal scaling vector  $\boldsymbol{\beta}(\omega)$  is not the same for all  $\omega$ .

We prefer to limit  $\boldsymbol{\beta}$  to the constraint that  $\boldsymbol{\beta}^* \mathbf{V}_\mathbf{X} \boldsymbol{\beta} = 1$ . Therefore the scaled series  $g(t, \boldsymbol{\beta})$  is unit variance. This will make the calculated spectral envelope more interpretable and make the magnitude of the elements of  $\boldsymbol{\beta}(\omega)$  easily comparable. Accordingly, the quantity  $\lambda(\omega)$  represents the largest power(variance) that can be obtained at the frequency  $\omega$  for any scaled series with unit variance.

With the optimal scaling vector  $\boldsymbol{\beta}(\omega)$ , equation (1) can be rewritten as:

$$\lambda(\omega) \mathbf{V}_\mathbf{X} \boldsymbol{\beta}(\omega) = \mathbf{P}_\mathbf{X}(\omega) \boldsymbol{\beta}(\omega) \quad (3)$$

It follows that  $\lambda(\omega)$  is the largest eigenvalue associated with the determinant equation:

$$|\mathbf{P}_\mathbf{X}(\omega) - \lambda(\omega) \mathbf{V}_\mathbf{X}| = 0 \quad (4)$$

$\boldsymbol{\beta}(\omega)$  is the corresponding eigenvector satisfying equation (3).

The viability of the spectral envelope for detecting common oscillation(s) comes from the following simple example. If the univariate time series  $x_i(t)$ ,  $1 \leq i \leq m$ ,  $-\infty < t < \infty$ , is in form of ‘common signal plus independent white noise’, say  $x_i(t) = s(t) + \varepsilon_i(t)$ , then in terms of power spectra,  $P_{x_i}(\omega) = P_s(\omega) + \sigma_\varepsilon^2$ , where  $var(\varepsilon_i) = \sigma_\varepsilon^2$ . A simple linear combination of  $x_i(t)$ , say  $\bar{x}(t) = m^{-1} \sum_{i=1}^m x_i(t)$ , will have as its power spectrum,  $P_{\bar{x}}(\omega) = P_s(\omega) + m^{-1} \sigma_\varepsilon^2$ . The signal to noise ratio of  $\bar{x}(t)$  has increased by a factor of  $m$  over the individual  $x_i(t)$ . Therefore, this indicates that the right linear combination of the original time series will enhance the signal and attenuate the noise. The spectral envelope method actually selects the optimal linear combination that can enhance the signal spectra and dampen the noise spectra at each frequency  $\omega$  for  $-1/2 < \omega \leq 1/2$ . This feature makes the spectral envelope particularly suitable for analyzing noise corrupted data (see simulation example in Section 2.4).

## 2.2 Simplified Definition of the Spectral Envelope

Denote  $\mathbf{V} = \text{diag}(\mathbf{V}_{\mathbf{X}})$ . The diagonal elements of  $\mathbf{V}$  and  $\mathbf{V}_{\mathbf{X}}$  are the same, but the off-diagonal elements of  $\mathbf{V}$  are zero. We can use  $\mathbf{V}$  instead of  $\mathbf{V}_{\mathbf{X}}$  in equation (1) and have a new expression for  $\lambda(\omega)$  (Stoffer 1999) :

$$\lambda(\omega) = \sup_{\boldsymbol{\beta} \neq 0} \left\{ \frac{\boldsymbol{\beta}^* \mathbf{P}_{\mathbf{X}}(\omega) \boldsymbol{\beta}}{\boldsymbol{\beta}^* \mathbf{V} \boldsymbol{\beta}} \right\} \quad (5)$$

The resulting  $\lambda(\omega)$  and  $\boldsymbol{\beta}(\omega)$  are slightly different than those in the equation (1). Under the condition that the rows of  $\mathbf{X}$  are mutually independent,  $\mathbf{V}$  is equal to  $\mathbf{V}_{\mathbf{X}}$  and equation (5) is the same as equation (1).

We also prefer to limit  $\boldsymbol{\beta}$  to the constraint such that  $\boldsymbol{\beta}^* \mathbf{V} \boldsymbol{\beta} = 1$ . As mentioned in Section 2.1, the idea in using equation (5) is that the right linear combination of the original time series will enhance the signal and dampen the noise.

## 2.3 Estimation of the Spectral Envelope

In practice, we can only have finite number of samples. So here we assume  $\mathbf{X} = [\mathbf{x}(0), \mathbf{x}(1), \dots, \mathbf{x}(n-1)] \in \mathfrak{R}^{m \times n}$  is a  $m \times n$  observed data matrix of  $m$  variables and  $n$  samples for each variable. Instead of applying spectral envelope method to the original data set, we prefer to work with the normalized data set which means that each variable is zero-mean with unit variance. The reasons for using normalized data are: (1) simplify the calculation; (2) make the calculated spectral envelope more interpretable; (3) make the calculated spectral envelope from different data sets comparable; (4) make the magnitude of the elements of  $\boldsymbol{\beta}(\omega)$  more meaningful and comparable. Here we assume  $\mathbf{X}$  is already an  $m \times n$  normalized data matrix.

With the normalized observed data  $\mathbf{X}$ , we can calculate its periodogram as:

$$\hat{\mathbf{I}}_n(\omega) = \frac{1}{n} \left[ \sum_{t=0}^{n-1} \mathbf{x}(t) \exp(-2\pi i t \omega) \right] \left[ \sum_{t=0}^{n-1} \mathbf{x}(t) \exp(-2\pi i t \omega) \right]^*, \quad -1/2 < \omega \leq 1/2 \quad (6)$$

The periodogram (equation (6)) provides a simple estimate of  $\mathbf{P}_{\mathbf{X}}(\omega)$ . But, the  $\hat{\mathbf{I}}_n(\omega)$  expression in equation (6) is still a continuous function of  $\omega$  which is not feasible. In practice, if  $n$  is a large integer, the fast Fourier transformation

provides for fast calculation of  $\hat{\mathbf{I}}_n(k/n)$ , for  $k = 1, \dots, [n/2]$ , where  $[n/2]$  is the greatest integer less than or equal to  $n/2$ . The frequencies  $\omega_k = k/n$ , for  $k = 1, \dots, [n/2]$ , are called the Fourier frequencies. Then the  $\hat{\mathbf{I}}_n(\omega_k)$  at these frequencies can be estimated as:

$$\hat{\mathbf{I}}_n(\omega_k) = \frac{1}{n} \left[ \sum_{t=0}^{n-1} \mathbf{x}(t) \exp(-2\pi it\omega_k) \right] \left[ \sum_{t=0}^{n-1} \mathbf{x}(t) \exp(-2\pi it\omega_k) \right]^* \quad (7)$$

The main drawbacks of using the periodogram directly is that it is not a consistent estimate of the PSD matrix. The confidence interval of the estimation is usually large (Shumway 1988). To overcome this problem, a smoothed periodogram estimate or a consistent spectral window estimate for  $\mathbf{P}_{\mathbf{X}}(\omega)$  can be used. The theory for estimating the PSD matrix of a vector process is well established (Brillinger 1981, Hannan 1970) and we discuss this briefly. One technique for smoothing is to take a symmetric moving average of the periodogram, that is:

$$\hat{\mathbf{P}}_{\mathbf{X}}(\omega_k) = \sum_{j=-r}^r h_j \hat{\mathbf{I}}_n(\omega_{k+j}) \quad (8)$$

where  $\{h_j\}$  are symmetric ( $h_j = h_{-j}$ ) positive weights and  $\sum_{j=-r}^r h_j = 1$ . The number  $r$  is chosen to obtain a desired degree of smoothness. Larger values of  $r$  leads to smoother estimates. However, one should be careful not to smooth away significant peaks. For  $\hat{\mathbf{P}}_{\mathbf{X}}(\omega_k)$  to be a consistent estimate, the weights must satisfy  $\sum h_j^2 \rightarrow 0$  as  $r \rightarrow \infty$ , but  $r/n \rightarrow 0$  as  $n \rightarrow \infty$ .

**Remark 1** *The optimal design of  $r$  and  $h_j$  is beyond the scope of this paper. For more detail, readers can refer to Stoffer et al. (2000). Throughout this paper,  $h_j$  is chosen as  $h_j = (r - |j| + 1)/(r + 1)^2$ , for  $|j| = 1, 2, \dots, r$ . In the illustrative example and the industrial case studies to follow,  $r$  was selected as an integer as 1 or 2.*

### 2.3.1 Estimation of $\lambda(\omega)$ using $\mathbf{V}_{\mathbf{X}}$

Using the estimated  $\hat{\mathbf{P}}_{\mathbf{X}}(\omega_k)$  and  $\hat{\mathbf{V}}_{\mathbf{X}} = \text{Cov}(\mathbf{X})$ , we can rewrite equation (1) as:

$$\hat{\lambda}(\omega_k) = \sup_{\boldsymbol{\beta} \neq 0} \left\{ \frac{\boldsymbol{\beta}^* \hat{\mathbf{P}}_{\mathbf{X}}(\omega_k) \boldsymbol{\beta}}{\boldsymbol{\beta}^* \hat{\mathbf{V}}_{\mathbf{X}} \boldsymbol{\beta}} \right\} \quad (9)$$

We denote the estimated  $\hat{\boldsymbol{\beta}}(\omega)$  at the Fourier frequencies as  $\hat{\boldsymbol{\beta}}(\omega_k)$ , and limit

$\hat{\boldsymbol{\beta}}(\omega_k)$  to the constraint that  $\hat{\boldsymbol{\beta}}(\omega_k)^* \hat{\mathbf{V}}_{\mathbf{X}} \hat{\boldsymbol{\beta}}(\omega_k) = 1$ .

For simplicity and without loss of generality,  $\hat{\lambda}(\omega_k)$  is defined to be the largest eigenvalue of  $\hat{\mathbf{H}}(\omega_k)$  where

$$\hat{\mathbf{H}}(\omega_k) = \hat{\mathbf{V}}_{\mathbf{X}}^{-\frac{1}{2}} \hat{\mathbf{P}}_{\mathbf{X}}(\omega_k) \hat{\mathbf{V}}_{\mathbf{X}}^{-\frac{1}{2}} \quad (10)$$

If we denote  $\hat{\boldsymbol{\beta}}_0(\omega_k)$  as the eigenvector of  $\hat{\mathbf{H}}(\omega_k)$  associated with  $\hat{\lambda}(\omega_k)$ , then the optimal scaling vector  $\hat{\boldsymbol{\beta}}(\omega_k)$  is defined by  $\hat{\boldsymbol{\beta}}(\omega_k) = \hat{\mathbf{V}}_{\mathbf{X}}^{-\frac{1}{2}} \hat{\boldsymbol{\beta}}_0(\omega_k)$ .

### 2.3.2 Estimation of $\lambda(\omega)$ using $\mathbf{V}$

Using the estimated  $\hat{\mathbf{P}}_{\mathbf{X}}(\omega_k)$  and  $\hat{\mathbf{V}} = \text{diag}(\hat{\mathbf{V}}_{\mathbf{X}}) \in \Re^{m \times m}$ , we can rewrite equation (5) as:

$$\hat{\lambda}(\omega_k) = \sup_{\boldsymbol{\beta} \neq 0} \left\{ \frac{\boldsymbol{\beta}^* \hat{\mathbf{P}}_{\mathbf{X}}(\omega_k) \boldsymbol{\beta}}{\boldsymbol{\beta}^* \hat{\mathbf{V}} \boldsymbol{\beta}} \right\} \quad (11)$$

Since the data has been normalized, we have  $\hat{\mathbf{V}} = \text{diag}\{1, 1, \dots, 1\} \in \Re^{m \times m}$ . Thus the constraint  $\hat{\boldsymbol{\beta}}(\omega_k)^* \hat{\mathbf{V}} \hat{\boldsymbol{\beta}}(\omega_k) = 1$  can be simplified as  $\hat{\boldsymbol{\beta}}(\omega_k)^* \hat{\boldsymbol{\beta}}(\omega_k) = 1$ . Then  $\hat{\lambda}(\omega_k)$  is the largest eigenvalue of  $\hat{\mathbf{P}}_{\mathbf{X}}(\omega_k)$ , and  $\hat{\boldsymbol{\beta}}(\omega_k)$  is the corresponding eigenvector.

## 2.4 Simulation Example

The following simulation example demonstrates the superiority of the performance of the spectral envelope method over the power spectrum and the SPCA method in detecting oscillation(s) and variable categorization.

### 2.4.1 Time Series Generation

This example consists of 12 time series generated with various sinusoid oscillations. In these time series,  $\varepsilon(t)$  is a white noise sequence with unit variance and  $t = 1, \dots, 512$ .

The first four time series are corrupted by moving average type colored noise and have base oscillation at frequency  $\omega_1 = 0.1Hz$ :

$$\begin{aligned}
x_1(t) &= 0.8 \cos(2\pi\omega_1 t) + \varepsilon(t) + 0.5\varepsilon(t-1) \\
x_2(t) &= 0.6 \cos[2\pi\omega_1(t-5)] + \varepsilon(t) + 0.5\varepsilon(t-1) \\
x_3(t) &= 0.4 \cos[2\pi\omega_1(t-15)] + \varepsilon(t) + 0.5\varepsilon(t-1) \\
x_4(t) &= 0.2 \cos[2\pi\omega_1(t-2)] + \varepsilon(t) + 0.5\varepsilon(t-1)
\end{aligned}$$

The next four time series are also corrupted by colored noise and have base oscillation at frequency  $\omega_2 = 0.3Hz$ :

$$\begin{aligned}
x_5(t) &= 0.9 \cos(2\pi\omega_2 t) + \varepsilon(t) - 0.5\varepsilon(t-1) \\
x_6(t) &= 0.7 \cos[2\pi\omega_2(t-7)] + \varepsilon(t) - 0.5\varepsilon(t-1) \\
x_7(t) &= 0.5 \cos[2\pi\omega_2(t-10)] + \varepsilon(t) - 0.5\varepsilon(t-1) \\
x_8(t) &= 0.3 \cos[2\pi\omega_2(t-20)] + \varepsilon(t) - 0.5\varepsilon(t-1)
\end{aligned}$$

The next two time series have oscillations at both frequencies  $\omega_1 = 0.1Hz$  and  $\omega_2 = 0.3Hz$ :

$$\begin{aligned}
x_9(t) &= 0.4 \cos[2\pi\omega_1(t-6)] + 0.5 \cos[2\pi\omega_2(t-8)] + \varepsilon(t) + 0.5\varepsilon(t-1) \\
x_{10}(t) &= 0.8 \cos[2\pi\omega_1(t-16)] + 0.6 \cos[2\pi\omega_2(t-4)] + \varepsilon(t) - 0.5\varepsilon(t-1)
\end{aligned}$$

The last two time series are moving average noise sequences:

$$\begin{aligned}
x_{11}(t) &= \varepsilon(t) + 0.5\varepsilon(t-1) \\
x_{12}(t) &= \varepsilon(t) - 0.5\varepsilon(t-1)
\end{aligned}$$

Before doing further analysis, all the time series are normalized to be zero-mean and unit variance. Figure 1 shows the time trends and the corresponding power spectra of these signals. The power spectra of the time series provide some information about the oscillations: For example, time series 1, 2 and 10 have peaks at frequency  $\omega_1 = 0.1Hz$ , and time series 5, 6 and 7 have peaks at frequency  $\omega_1 = 0.3Hz$ . However, for other time series, the power spectra do not provide clear indication of the underlying oscillations.

#### 2.4.2 SPCA Analysis

Figure 2 shows the plot of the first two principle components (PCs) of the spectral analysis of the signals. These two PCs explain over 83% of the variability of the spectra. These two PCs give indication of the base oscillations

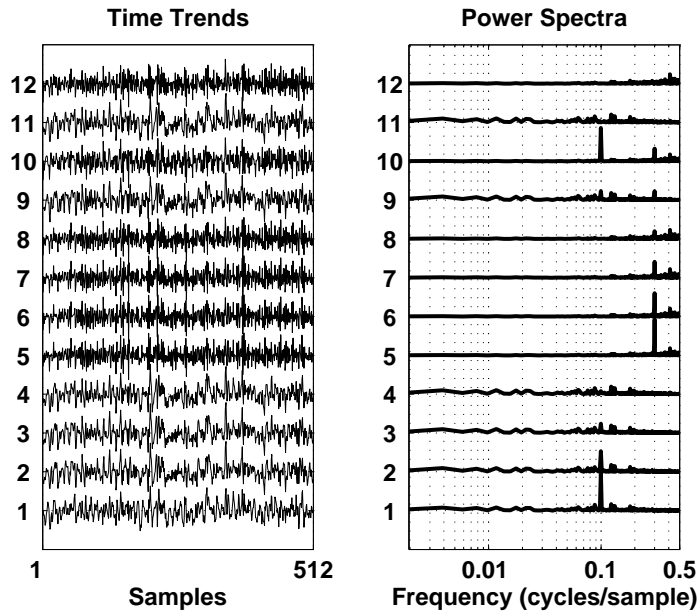


Fig. 1. Time trends and power spectral of the 12 time series

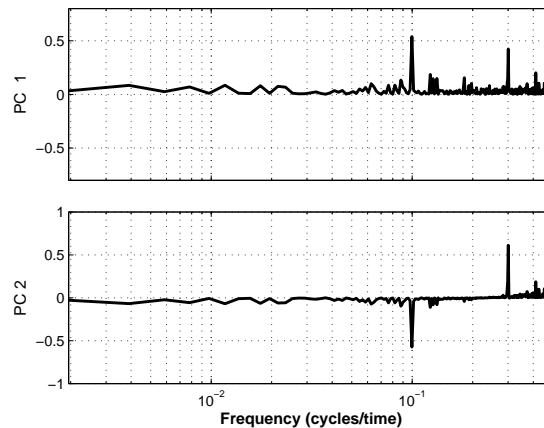


Fig. 2. SPCA PCs plot of the 12 time series

at  $0.1Hz$  or  $0.3Hz$ . However, these two PCs also indicate misleading peaks at frequencies  $0.12Hz$  and  $0.4Hz$ , which actually are not oscillation frequencies in this example. As can be noticed in Figure 2, part of PC2 contains a negative spectrum, which is not causal since no signal has negative power. The same problem also occurs in the industrial case studies analyzed in Section 6 and 7. Thus the physical interpretation of SPCA analysis is difficult in such cases. A modified extension of SPCA based on non-negative matrix factorization that conforms to the causality of positive power spectral is under development by (Tangirala and Shah 2005). Figure 3 shows the two-dimensional scores plot from SPCA analysis of the 12 time series. Each tag maps to a point in the two-dimensional space. There is no obvious cluster. From examination of Fig-

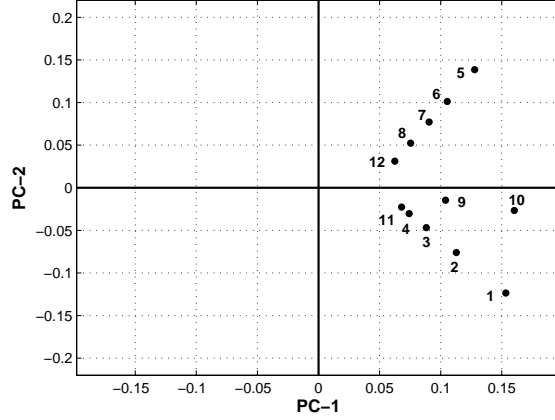


Fig. 3. SPCA scores plot of the 12 time series

ure 2 and 3 one could say: time series 1, 2 and 3 have positive PC1 value and negative PC2 value, thus they probably have oscillation at  $0.1Hz$ ; time series 5, 6, 7 and 8 have both positive PC1 value and PC2 value, then they probably have oscillation at  $0.3Hz$ . But for the rest time series, it is hard to analyze their oscillation frequencies: the multiple oscillations in time series 9 and 10 are not detected; time series 11 and 12 can not be identified as not having oscillation.

#### 2.4.3 Oscillation Detection via the Spectral Envelope Method

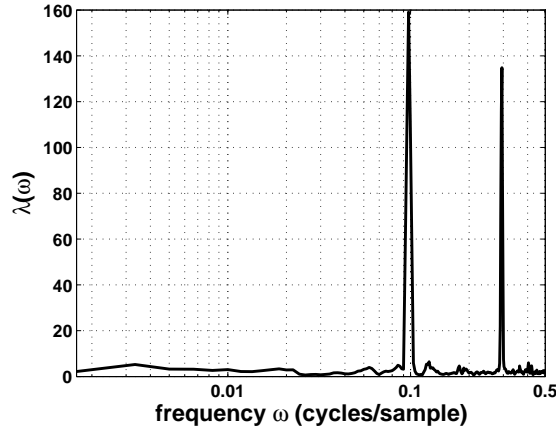


Fig. 4. Spectral Envelope of the 12 time series

Figure 4 shows the spectral envelope calculated using equation (9). In this analysis, we have  $n = 512$ ,  $m = 12$  with weights  $\{h_0 = 3/9, h_{\pm 1} = 2/9, h_{\pm 2} = 1/9\}$ . There are two significant peaks at frequencies:  $51/512 \approx 0.1Hz$  and  $154/512 \approx 0.3Hz$ , which means that the scaled series could have significantly more power at these two frequencies than at any other frequencies. It further implies that some of (or all of) the 12 time series may have significant power

at  $0.1Hz$  and/or  $0.3Hz$ . There are no misleading peaks at other frequencies. Therefore, it can be concluded that the spectral envelope method can clearly detect the multiple oscillations present in the time series. Compared with SPCA, the spectral envelope method is frequency sensitive and thus is better in detection of multiple oscillations; also this method can attenuate the noise spectra and thus is better in analyzing noise corrupted data. Variable categorization for this simulation example is discussed in next section.

**Remark 2** *Another oscillation detection technique is the auto-covariance functions (ACF) based method (Thornhill et al. 2003a). It uses the zero crossings of filtered auto-covariance functions to detect and categorize oscillations. Compared with the spectral envelope method, this method has some limitations in application:*

- *This method is prone to false detections because the algorithm uses ideal band pass filters. If it is a narrow band pass filter, the filtered data may be oscillatory and the algorithm may provide misleading results. This difficulty can be partly overcome by using other advanced filtering technique, but this will increase the complexity of the algorithm.*
- *The presence of noise and multiple oscillations may destroy the regularity of the zero crossings of the ACF. In this case, this method may detect none or only one oscillation, though the spectrum may show multiple distinct peaks.*
- *This method needs a minimum of five cycles in the auto-covariance function to detect oscillation, which is often hard to obtain, particularly in the case of a long oscillation.*

### 3 Variable Categorization

After detecting the oscillation(s), the next task is to group the variables that oscillate at a common frequency. The ACF based method can be used to categorize similar oscillations, but this method has some limitations in application as mentioned earlier. The SPCA can also categorize variables using the scores plot. However, there are some limitations of applying scores plot as well:

- If the selected PCs can not detect the oscillation frequencies, then the scores plot may not always deliver any useful information.
- In practice, due to noisy data, the clustering in the scores plot is not very obvious. (See the two industrial case studies in Section 6 and Section 7.)
- One variable can belong to one cluster only in the scores plot. Thus the variables that have multiple oscillations are usually hard to analyze using the scores plot. The scores plot may only capture one main oscillation and lose information at the other frequencies. Therefore it is particular difficult to categorize all variables that have common multiple oscillations

- The visualization of the scores plot is limited by the number of PCs. A scores plot in more than 3-D is hard to analyze.

To overcome the disadvantages of ACF and SPCA methods, we utilize the optimal scaling vector  $\hat{\beta}(\omega_k)$  to categorize the variables having common oscillation(s). The magnitude of the optimal scalings (elements of the optimal scaling vector  $\hat{\beta}(\omega_k)$ ) is a measure of the contribution of each time series to the spectral envelope  $\lambda(\omega_k)$  at frequency  $\omega_k$ . Once a certain oscillation frequency is identified, then one can investigate the magnitude of the optimal scalings at that frequency. The time series having large optimal scaling magnitude are the ones that contribute the most to the spectral envelope, and thus are the ones having oscillation at that frequency. Furthermore, statistical hypothesis test can be performed to check whether or not a particular element of  $\hat{\beta}(\omega_k)$  is zero. This will greatly help us to identify the variables that have oscillations.

In this section, we only use the optimal scaling vector  $\hat{\beta}(\omega_k)$  calculated by equation (11). The reason why equation (11) is preferred to equation (9) is that equation (11) leads to simple calculation and one can avoid calculating equation (10) which involves the computation of  $\hat{\mathbf{V}}_{\mathbf{X}}^{-\frac{1}{2}}$ .

### 3.1 Statistical Hypothesis Test on $\hat{\beta}(\omega)$

Details on distribution of the sample spectral envelope and the optimal scalings can be found in Stoffer *et al.* (1993) and McDougall *et al.* (1997). Here we briefly summarize the main result: if  $\hat{\mathbf{P}}_{\mathbf{X}}(\omega)$  is a consistent estimator and if  $\lambda(\omega)$  is a distinct eigenvalue, then  $\nu_n[\hat{\beta}(\omega) - \beta(\omega)]$  converges ( $n \rightarrow \infty$ ) to a complex multivariate normal distribution. The term  $\nu_n$  depends on the type of estimator being used. For example, if a weighted average as in equation (8) is used,  $\nu_n = (\sum_{j=-r}^r h_j^2)^{1/2}$ . The asymptotic ( $n, r \rightarrow \infty$ ) covariance matrix of the sample optimal scaling vector  $\hat{\beta}(\omega)$ , say  $\mathbf{V}_{\beta}(\omega)$ , is given by:

$$\mathbf{V}_{\beta}(\omega) = \nu_n^{-2} \lambda_1(\omega) \sum_{l=2}^m \lambda_l(\omega) [\lambda_1(\omega) - \lambda_l(\omega)]^{-2} \beta_l(\omega) \beta_l^*(\omega) \quad (12)$$

where  $\{\lambda_1(\omega) = \lambda(\omega), \lambda_2(\omega), \dots, \lambda_m(\omega)\}$  are the eigenvalues of  $\mathbf{P}_{\mathbf{X}}(\omega)$  arranged in decreasing order, and  $\{\beta_1(\omega) = \beta(\omega), \beta_2(\omega), \dots, \beta_m(\omega)\}$  are the corresponding eigenvectors.

In addition, the distribution

$$\frac{2|\hat{\beta}_{1,j}(\omega) - \beta_{1,j}(\omega)|^2}{\sigma_j(\omega)} \quad (13)$$

is approximately a Chi-square distribution with 2 degrees of freedom, where  $\hat{\beta}_{1,j}(\omega)$ ,  $j = 1, \dots, m$  is the  $j$ th element of the optimal scaling vector, and  $\sigma_j(\omega)$  is the  $j$ th diagonal element of  $\hat{\mathbf{V}}_\beta(\omega)$ , the estimate of  $\mathbf{V}_\beta(\omega)$ . One could use equation (12) and (13) to form confidence regions for each  $\hat{\beta}_{1,j}(\omega)$ . Also, one can check whether or not  $\hat{\beta}_{1,j}(\omega)$  is zero by comparing  $2|\hat{\beta}_{1,j}(\omega)|^2/\sigma_j(\omega)$  with  $\chi_2^2(\alpha)$ , the  $\alpha$  upper tail cutoff of the Chi-square distribution. If  $2|\hat{\beta}_{1,j}(\omega)|^2/\sigma_j(\omega) > \chi_2^2(\alpha)$ , then the null hypothesis ' $\beta_{1,j}(\omega) = 0$ ' is rejected with  $(1 - \alpha)$  confidence. Then one can conclude that, with  $(1 - \alpha)$  confidence, the corresponding time series does have oscillation at that frequency. If  $2|\hat{\beta}_{1,j}(\omega)|^2/\sigma_j(\omega) < \chi_2^2(\alpha)$ , then the null hypothesis ' $\beta_{1,j}(\omega) = 0$ ' is accepted, the corresponding time series can be treated as not having oscillation in the statistical sense. This statistical procedure is particularly useful in automating the task of finding common oscillations in a data set.

Table 1  
Variables categorization for simulation example

at 0.1Hz		at 0.3Hz	
Series	Test statistic	Series	Test statistic
1	144	1	5
2	187	2	5
3	71	3	5
4	16	4	5
5	0	5	1289
6	0	6	3920
7	0	7	1008
8	0	8	214
9	66	9	130
10	1392	10	263
11	0	11	5
12	0	12	6

Table 1 presents the optimal scalings and appropriate test statistics (based on equation. (13)) at the oscillation frequencies for each time series of the example in Section 2.4. The smoothing weights used are the same  $\{h_0 = 3/9, h_{\pm 1} = 2/9, h_{\pm 2} = 1/9\}$ . If we choose  $\chi_2^2(0.001) = 13.82$ , then the conclusion reached from the hypothesis test is that the test has successfully identified the correct oscillation variables at both 0.1Hz and 0.3Hz with a 99.9% confidence, i.e. variables 1 - 4, 9 and 10 oscillate at 0.1Hz frequency and variables 5 - 10 oscillate at 0.3Hz frequency. From the intersection of these two sets, it is clear that variables 9 and 10 have oscillations at frequencies of 0.1Hz and 0.3Hz.

## 4 Root Cause Diagnosis

Root cause diagnosis is a challenging problem in the area of detection and diagnosis of plant-wide oscillations. Due to the complexity of large-scale plants and the difficulty of determining cause-effect relationship, it is difficult to conclude whether a certain variable is the root cause just simply from the analysis of plant data. Process knowledge and plant test are indispensable in determining the real root cause. The contribution of current data based root cause diagnosis techniques (Thornhill *et al.* 2001, Thornhill *et al.* 2003b) is to isolate the few key variables as the root cause candidates, or at least identify those variables that are ‘physically’ close to the root cause. This will reduce the workload and cost of further plant tests to determine the real root cause. The idea in Thornhill *et al.* (2001) or Thornhill *et al.* (2003b) is to use distortion factor or non-linearity index as signatures to isolate the key oscillation variables as the root cause candidates. In this section, we propose a new index called the oscillation contribution index (*OCI*) to serve the same purpose.

The main idea of *OCI* is to utilize the optimal scalings of the oscillation variables that are identified by the statistical hypothesis test in last section. The root cause(s) is most likely to be within these variables that have oscillations. If  $x_j(t)$  is one of the oscillation variables, then its *OCI* is defined to be:

$$OCI_j(\omega) = \frac{\hat{\beta}_{1,j}(\omega)}{2\sigma_{\hat{\beta}}(\omega)} \quad (14)$$

where  $\sigma_{\hat{\beta}}(\omega)$  is the standard deviation of the optimal scalings of all the identified variables that have oscillations. The *OCI* is an indicator of the contribution of each oscillation variables to the spectral envelope peak at the oscillation frequency. We use the *OCI* to isolate the key variables as the root cause candidates. A general criteria is that the variables having  $OCI(\omega) > 1$  are the likely root cause variables at frequency  $\omega$  because they contribute most to the spectral envelope at the oscillation frequency. The test can be more stringent and discriminating if a three-sigma test is used. The industrial case studies in later sections will demonstrate the efficacy of using the *OCI* to isolate the key variables.

**Remark 3** *If there is no root cause, as is the case in the previous generated time series example, or the root cause variable is not included in the original data set, then of course the OCI would not be able to find it. The calculated OCI in this situation can be used to rank the variables according to their contribution to the spectral envelope at the oscillation frequency and find out the variables that have significant oscillations. In such cases the OCI for each*

*variable can simply be interpreted as the signal strength by that variable towards the frequency of interest. In general the purpose of the OCI is to provide a ranked list of variables that are the likely root cause variables or lie physically close to the root cause. This ranked list combined with process flowsheet, loop configuration and other pertinent information can lead to the correct diagnosis of root causes.*

## **5 Summary of the Procedure to Detect and Diagnose Plant-Wide Oscillations**

Based on the result in previous sections, the following procedure to detect and diagnose plant-wide oscillations is proposed:

- (1) Normalize the data matrix so that each variable has zero-mean and unit variance;
- (2) Calculate the spectral envelope using equation (1) or (5) to find out the main oscillation frequencies;
- (3) Do statistical hypothesis test to identify the variables that have oscillations at those oscillation frequencies found in step (2).
- (4) Use the *OCI* to isolate a few key variables as the root cause candidates. Examine these variables to confirm the cause of oscillations.

## **6 Industrial case study 1**

An industrial data set was provided by the Advanced Controls Technology group of Eastman Chemical Company. Figure 5 shows the process schematic of the plant, which contains three distillation columns, two decanters and several recycle streams. There are 15 control loops and 15 indicators on the schematic. The Advanced Controls Technology group had identified a need for diagnosis of a common disturbance with an oscillation period of about 2 hours. In this section, the proposed procedure is applied to this data set to demonstrate its efficacy in detection and diagnosis of the root cause of this oscillation.

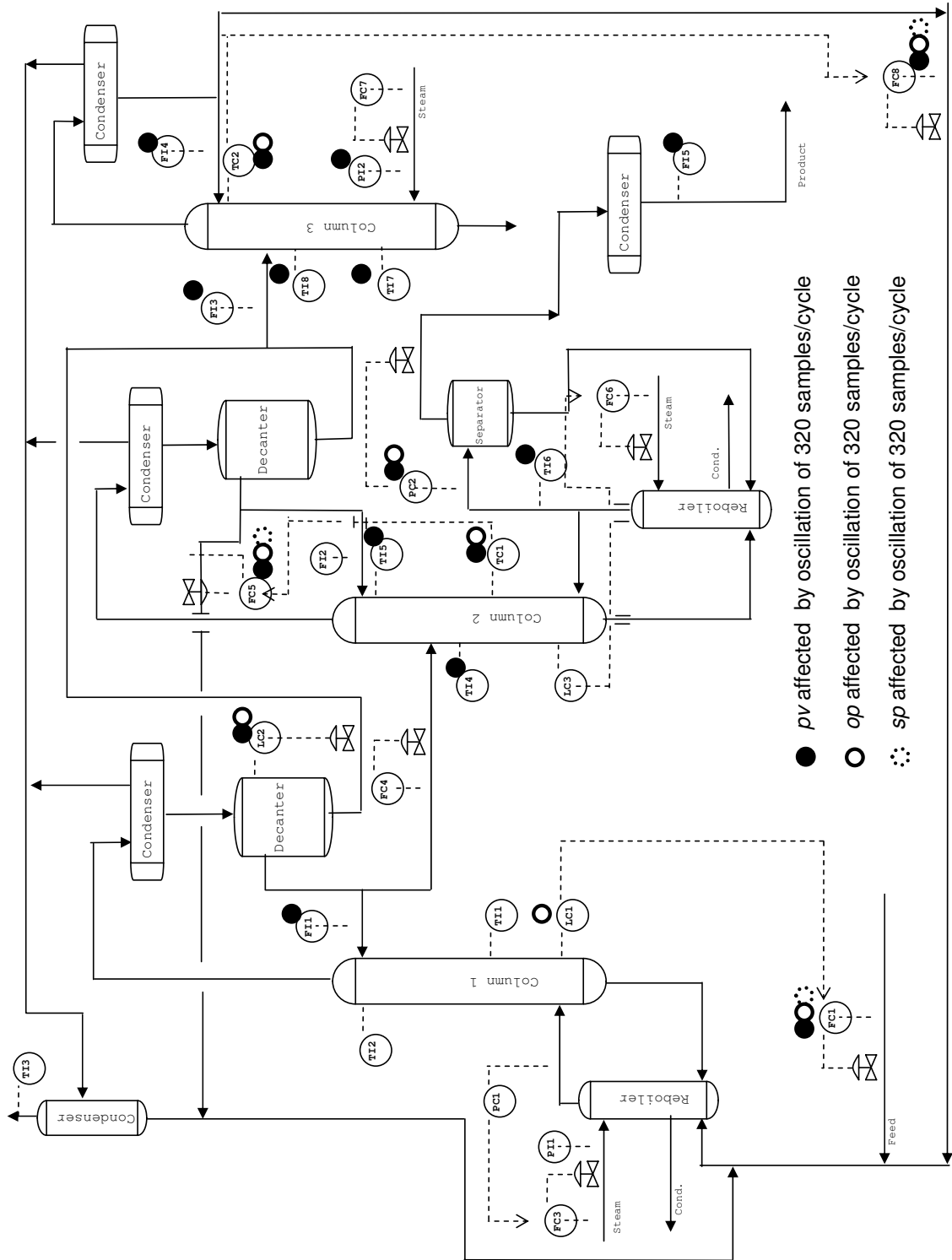


Fig. 5. Process schematic. The oscillation variables are marked by circle symbols.

## 6.1 Data Description

The provided data set contains 48 variables: 14 process variables (*pv*'s), 14 controller outputs (*op*'s), 15 indicator variables and 5 cascade loop setpoints (*sp*'s). The whole data set contains 96 hours of data with a sampling time of 20s. Thornhill *et al.* (2003b) used the second 48 hour data window to analyze the oscillations. Here, we use the first 48 hour data window where each variable has 8640 observations. In this case study (also for the case study in Section 7), AC, FC, LC, PC and TC represent composition, flow, level, pressure and temperature tags, respectively, that are controlled. Similarly, FI, LI, PI, TI and SI represent the flow, level, pressure, temperature and rotor speed tags, respectively, that are indicators only. We denote the set point, process value and the controller output as *sp*, *pv* and *op*, respectively.

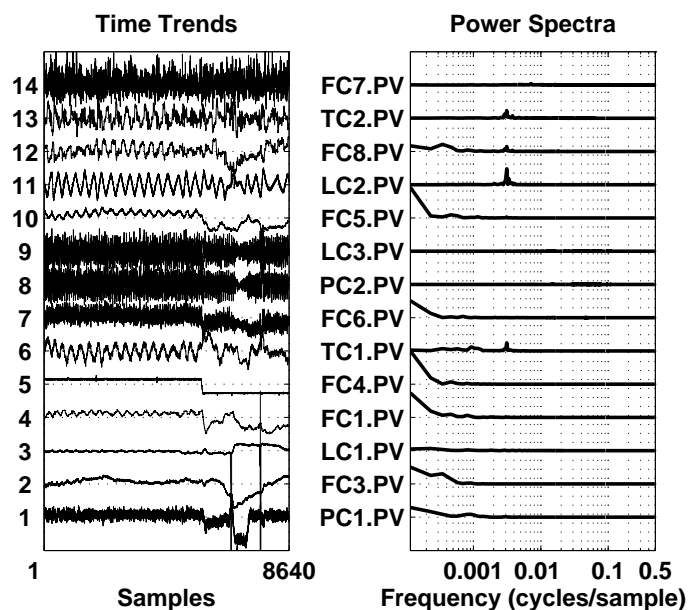


Fig. 6. Time trend and power spectra of 14 *pv*'s

Figure 6 shows the time trends and power spectra of the 14 *pv* variables. The power spectra indicate the presence of oscillation at the frequency of 0.003 cycles/sample (or about 333 samples/cycle, nearly a period of 2 hours). This oscillation had propagated through out the adjacent units and affected many variables in the process.

## 6.2 SPCA Analysis

Figure 7 shows the first two principle components (PCs) plot. These two PCs explain 86.26% variability of the spectra. The second PC has a peak near the

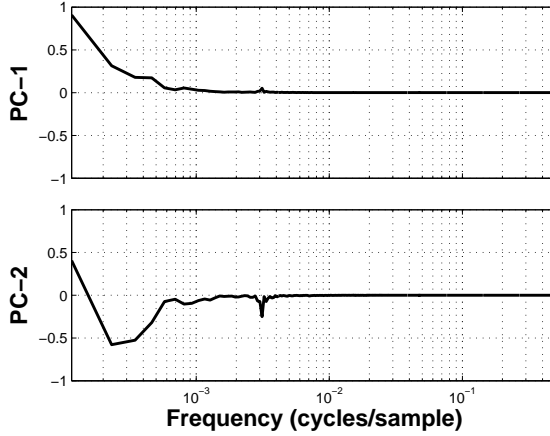


Fig. 7. SPCA PCs plot of the 48 variables

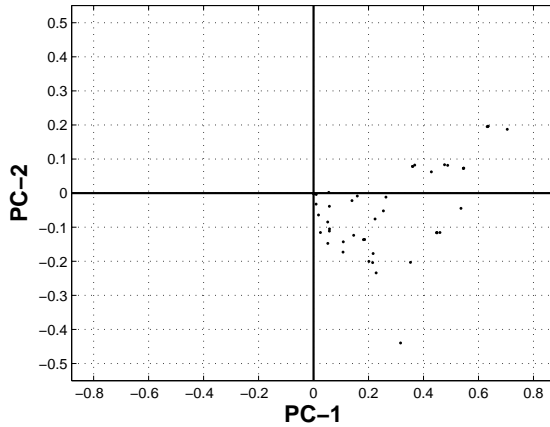


Fig. 8. SPCA scores plot of the 48 variables

frequency of 0.003 cycles/sample which indicates the oscillation of interest. However, the negative spectrum in PC2 is not causal from a physical point of view. As for the categorization, the two-dimensional scores plot (Figure 8) has no meaningful clustering. It is hard to analyze the frequency features of each variable using SPCA. Other PCs do not help the analysis either.

### 6.3 Analysis Steps in the Detection and Diagnosis of Oscillations

#### 6.3.1 Oscillation Detection

Figure 9 shows the spectral envelope (from equation (11)) of the 48 variables. This spectral envelope is estimated using triangular smoothing with  $r = 1$  and weights  $\{h_0 = 1/2, h_{\pm 1} = 1/4\}$ . In the spectral envelope, there are clear low frequency features. This is probably because the data is from a long term operation and there exists extremely long period influences such as diurnal weather

effects that impact the process. This low frequency feature is quite common in industrial routine operating data and it is also present in the industrial case study (2) in Section 7. In addition to this low frequency feature, there is a clear peak at a frequency of  $27/8640 \approx 0.0031$  cycles/sample, indicating an oscillation with a period of 320 samples/cycle. This is exactly the oscillation that the Advanced Controls Technology group of Eastman Chemical Company wanted to detect and diagnose.

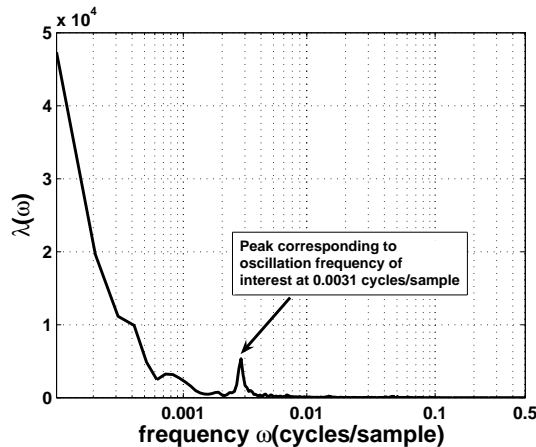


Fig. 9. Spectral Envelope of the 48 variables

### 6.3.2 Variable Categorization

Table 2

Variables categorization for industrial case study (1)

Tag No.	Test statistic	Tag No.	Test statistic	Tag No.	Test statistic
LC1.op	44	PC2.pv	54	FC8.op	789
FI1.pv	39	PC2.op	52	FI4.pv	143
FC1.sp	44	FC5.sp	395	TC2.pv	254
FC1.pv	46	FC5.pv	428	TC2.op	423
FC1.op	35	FC5.op	165	TI8.pv	188
TI5.pv	78	FI5.pv	56	TI7.pv	25
TI4.pv	1156	LC2.pv	526	PI2.pv	26
TC1.pv	1285	LC2.op	583	FI3.pv	364
TC1.op	397	FC8.sp	418		
TI6.pv	91	FC8.pv	419		

Table 2 shows the variables that have the test statistic value bigger than  $\chi_2^2(0.001) = 13.82$  at the oscillation frequency. In other words, we can confirm

with 99.9% confidence that these are the variables that have oscillations. These oscillation variables are also marked by dark circle symbols in Figure 5.

### 6.3.3 Oscillation Diagnosis

Table 3

Ranked list of variables having  $OCI$  bigger than 1 at the oscillation frequency

Tag No.	$OCI$	Tag No.	$OCI$
LC2.pv	2.09	FC8.pv	1.17
LC2.op	1.88	TC2.op	1.11
TI4.pv	1.51	FC8.sp	1.11
TC1.pv	1.48	TI5.pv	1.10
TC2.pv	1.44	TI8.pv	1.07

Table 3 shows the variables that have  $OCI$  bigger than 1 at the oscillation frequency. They are treated as the root cause candidates. Among all the variables, the  $pv$  and  $op$  of the level control loop LC2 have the largest  $OCI$  at the oscillation frequency. This result indicates that the LC2 loop contributes most to the spectral envelope at the oscillation frequency and we should take this loop as the first root cause candidate.

To confirm and isolate the root cause, the  $pv$  and  $op$  data for this loop was analyzed further. Higher order statistical method described in Choudhury *et al.* (2004b), Choudhury *et al.* (2004a) and Choudhury (2004) are used to investigate this problem.

In Choudhury *et al.* (2004b), two indices - the Non-Gaussianity Index ( $NGI$ ) and the Non-Linearity Index ( $NLI$ ) - have been defined based on bicoherence of a time series signal. When both  $NGI$  and  $NLI$  are greater than zero, the signal is described as non-Gaussian and nonlinear and it is inferred that the loop in question exhibits significant non-linearity. For a control loop, this test is applied to the error signal ( $sp - pv$ ) as the error signal is often more stationary than  $pv$  or  $op$  signals. Assuming that the process is linear and no nonlinear disturbances enter the loop, the nonlinearity can be attributed to the control valve. Once a nonlinearity is detected using higher order statistical method-based  $NGI$  and  $NLI$  indices, the  $pv-op$  plot is used to diagnose and isolate its cause. It is well known (Hägglund 1995, Rengaswamy *et al.* 2001, Choudhury *et al.* 2005a, Choudhury 2004) that the presence of stiction in control valve in a control loop produces limit cycles in the controlled variable ( $pv$ ) and the controller output ( $op$ ). For such a case, the  $pv-op$  plot shows elliptical cyclic patterns, which are taken as a signature of valve stiction. If no such patterns are observed, it is concluded that there may be valve problems other than stiction. Note that for the cases of tightly tuned controller or a process with

time delay, the  $pv-op$  plot may also exhibit elliptical patterns. But they do not add nonlinearity in a control loop. Therefore, these cases do not pass the nonlinearity test. The  $pv-op$  plot is investigated only after a successful nonlinearity detection in the loop. That is why the  $pv-op$  plot should not be used alone to detect stiction. This must be used in conjunction with the nonlinearity test. For a detailed discussion on these issues, refer to Choudhury *et al.* (2004b), Choudhury *et al.* (2006).

Figure 10 shows the diagnostic plots for this loop. Figure 10(a) shows that there are significantly large peaks in the bicoherence plot indicating a non-linear loop. The values of  $NGI$  and  $NLI$  for this loop are 0.15 and 0.42, respectively, which clearly indicates that the loop exhibits nonlinearity. Once a loop nonlinearity is detected, it should be checked whether this is due to stiction or other process nonlinearity. Figure 10(b) shows the  $pv-op$  plot for this loop. The plot clearly shows an elliptic pattern indicating the presence of stiction in the control valve. The apparent stiction is quantified to be approximately 3% using the method described in Choudhury *et al.* (2006).

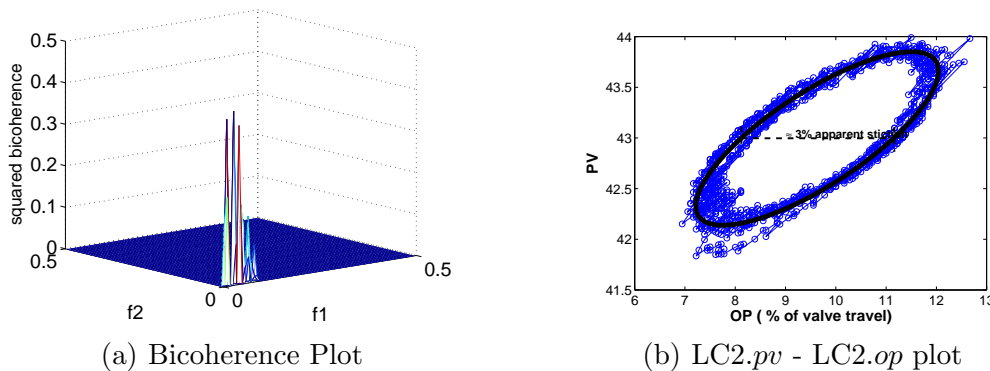


Fig. 10. Oscillation diagnosis plots for LC2 loop

Similar results of root cause diagnosis were also discussed in Thornhill *et al.* (2003b). It was reported that the control valve of loop LC2 suffered from a deadband problem (Thornhill *et al.* 2003b). It has been confirmed that the control valve caused control variable  $LC2.pv$  to oscillate, and the oscillation passed through the feedback controller and made the controller output  $LC.op$  also to oscillate. After that, the oscillations propagated to the temperature control loop TC1 in the second distillation column and caused the temperature to oscillate. This is the reason why temperature indicator  $TI4.pv$  and control variable  $TC1.pv$  had oscillations too. For more information, refer to Thornhill *et al.* (2003b).

## 7 Industrial case study 2

An industrial data set was provided by the courtesy of Mitsubishi Chemical Corporation (MCC), Mizushima, Japan. Figure 11 shows the process schematic of the plant. The plant personnel reported oscillations with a period of about 2 ~ 3 hours in the condenser level of a distillation column. These oscillations actually propagated through out the plant, causing sub-optimal operation and large economic losses. Previous attempts for oscillation detection and root cause diagnosis by considering only the level and variables directly affecting them were not successful. In this section, the newly proposed procedure is applied to this large data set to detect and diagnose the cause of these plant-wide oscillations.

### 7.1 Data Description

The provided data set consists of 58 variables: 27 process variables (*pv*'s), 15 controller outputs (*op*'s), 16 indicator variables. Each variable has 3600 observations with a sample interval of 1 minute, which corresponds to data over 2 days of operation.

### 7.2 SPCA Analysis

Figure 12 shows a plot of the first three principle components (PCs). These three PCs explain 94.68% variability of the spectra. The third PC has a peak near the frequency of 0.007 cycles/sample (or 143 samples/cycle) which indicates the oscillation of interest. However, the negative spectra in PC2 and PC3 is hard to interpret. As for the categorization, the three-dimensional scores plot (Figure 13) has no meaningful clustering. It is hard to analyze the frequency features of each variable. Other PCs do not help the categorization either.

### 7.3 Analysis Steps in the Detection and Diagnosis of Oscillations

#### 7.3.1 Oscillation Detection

Figure 14 shows the spectral envelope (from equation (11)) of the 58 variables. This spectral envelope is estimated using triangular smoothing with  $r = 1$  and weights  $\{h_0 = 1/2, h_{\pm 1} = 1/4\}$ . Besides the low frequency features, there is a clear peak at the frequency of  $25/3600 \approx 0.0069$  cycles/sample. This peak

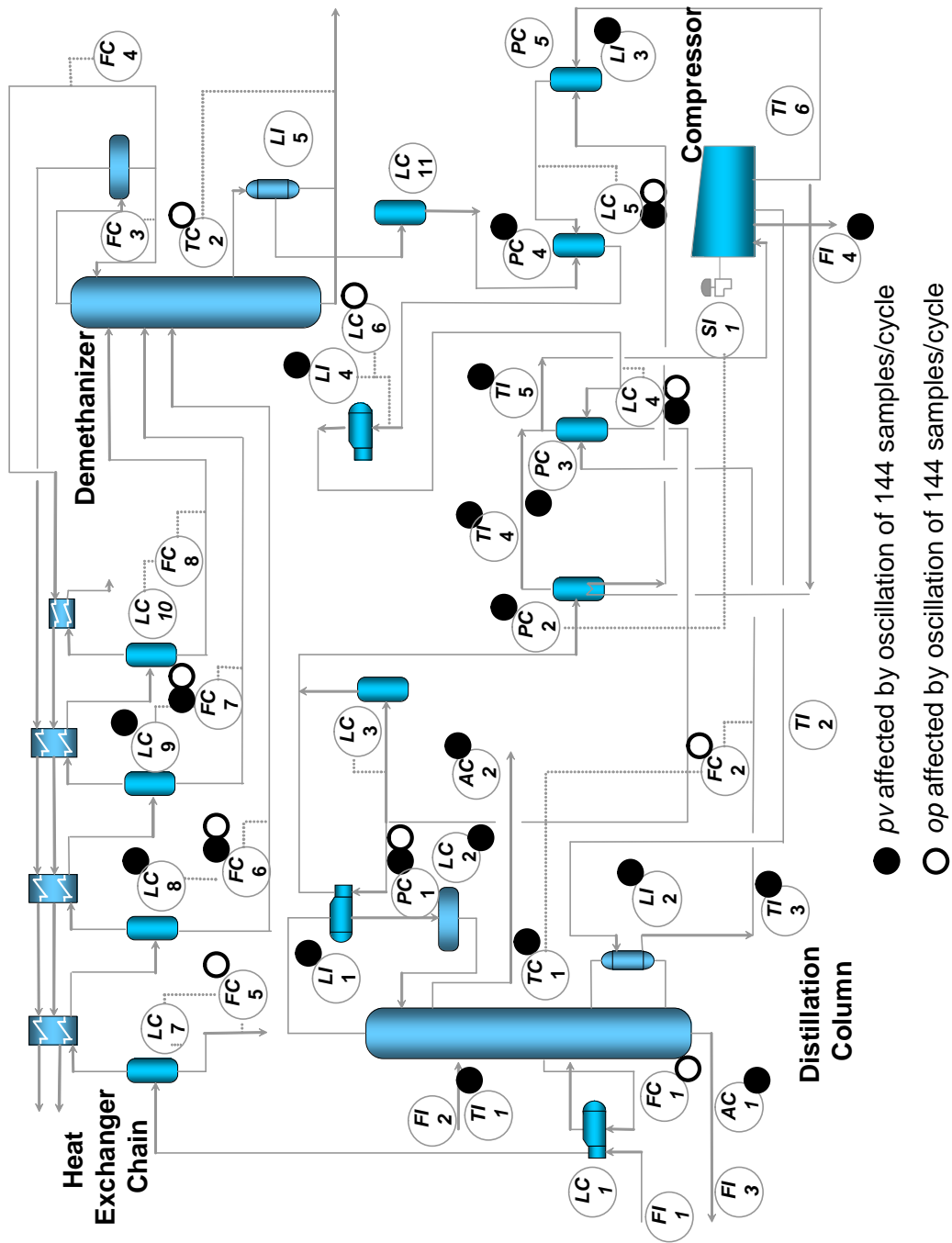


Fig. 11. Process schematic. The oscillation variables are marked by circle symbols.

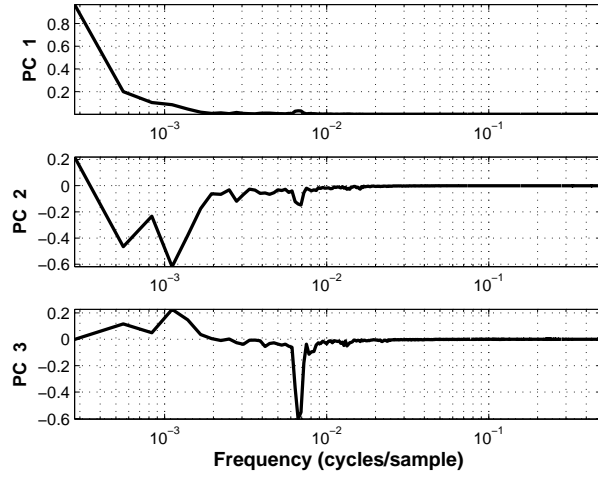


Fig. 12. SPCA PCs plot of the 58 variables

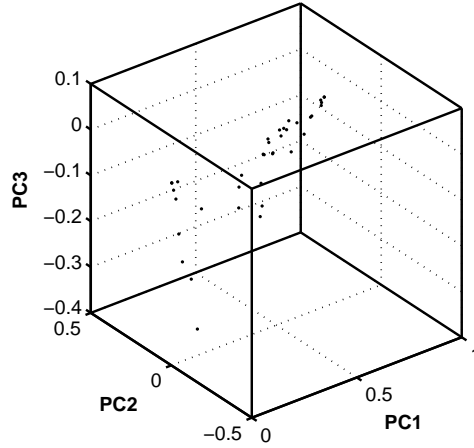


Fig. 13. SPCA scores plot of the 58 variables

indicates a oscillation with a period of 144 samples/cycle, which is exactly the oscillation that the plant personnel wanted to detect and diagnose.

### 7.3.2 Variable Categorization

Table 4 shows the variables that have the test statistic value bigger than  $\chi_2^2(0.001) = 13.82$  at the oscillation frequency. In other words, with a confidence of 99.9%, these are the variables that have oscillations. These oscillation variables are also marked by dark circle symbols in Figure 11.

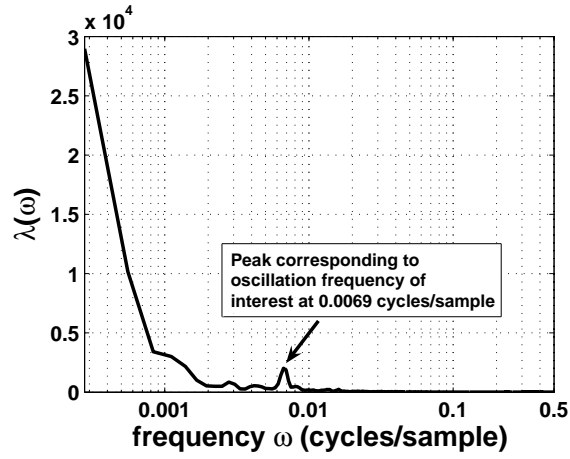


Fig. 14. Spectral Envelope of the 58 variables

Table 4  
Variables categorization for industrial case study (2)

Tag No.	Test statistic	Tag No.	Test statistic	Tag No.	Test statistic
TI1.pv	492	TI3.pv	96	FI4.pv	56
FC1.op	23	PC2.pv	269	LI4.pv	87
AC1.pv	33	TI4.pv	141	LC6.op	93
LI1.pv	78	PC3.pv	100	TC2.op	25
PC1.op	2867	TI5.pv	34	FC5.op	60
PC1.pv	3385	LC4.op	1310	LC8.pv	37
LC2.pv	398	LC4.pv	325	FC6.op	23
AC2.pv	120	PC4.pv	92	FC6.pv	32
TC1.pv	22	LC5.op	125	LC9.pv	36
LI2.pv	56	LC5.pv	965	FC7.op	21
FC2.op	61	LI3.pv	268	FC7.pv	40

Table 5  
Ranked list of variables having *OCI* bigger than 1 at the oscillation frequency

Tag No.	<i>OCI</i>	Tag No.	<i>OCI</i>
PC1.pv	2.23	LI3.pv	1.36
TI4.pv	1.68	LC4.pv	1.33
LC5.pv	1.64	LC8.pv	1.18
LI1.pv	1.48	LC5.op	1.08
LC4.op	1.38		

### 7.3.3 Oscillation Diagnosis

Table 5 shows variables that have  $OCI$  bigger than 1 at the oscillation frequency. They are treated as the root cause candidates. Among all the variables, variable  $PC1.pv$  has largest  $OCI$ . This result indicates that this particular (PC1) loop contributes most to the spectral envelope at the oscillation frequency and we should examine this loop as the first root cause candidate.

Again, we followed the same procedure as in the earlier industrial case study (1) to diagnose the cause of the oscillations. Figure 15(a) shows the bicoherence plot. The values of  $NGI$  and  $NLI$  are 0.18 and 0.54 respectively, indicating a nonlinear loop. The elliptical pattern in Figure 15(b) indicates that this control valve is suffering from stiction.

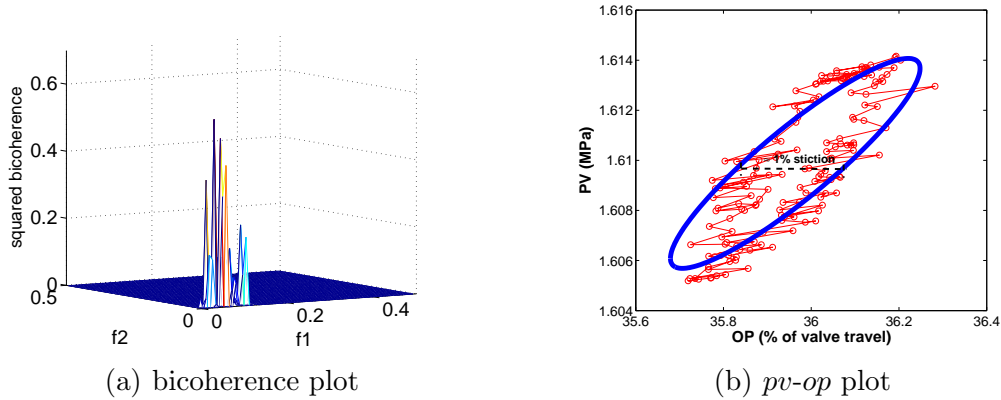


Fig. 15. Diagnostic plots for the PC1 loop

Further plant tests have confirmed that this loop with a sticky valve was indeed the leading cause of plant-wide oscillations. Before the ‘plant-shutdown’ for maintenance could proceed, a simple closed-loop test described in Choudhury *et al.* (2005b) was performed on loop PC1 to confirm the presence of valve stiction. This simple closed-loop test requires one to change the controller gain and observe whether there is a change in frequency of oscillation. If the frequency of oscillation changes, this indicates that the oscillation is generated within the loop and is not caused by an external disturbance. The results obtained from this test on loop PC1 confirmed that the oscillation was generated within the loop PC1. MCC engineers have now confirmed that the sticky valve in this loop plus interactions from other loops were the main causes of oscillations.

## 8 Concluding Remarks

In this paper, the concept of spectral envelope has been modified and extended so that it is simple to apply for detecting plant-wide oscillations. This method is good at detecting single or multiple oscillations. The oscillation detection can be carried out using routine operating data and the calculation of spectral envelope is straightforward. In comparison to the ACF based method, the spectral envelope method does not suffer any limitation on the minimum number of oscillation cycles and it does not require the design of any filter. It can detect all oscillations in a single step and therefore the potential for automating this method is significant. The method can also identify variables with multiple oscillations.

Further more, statistical hypothesis test can be performed to identify variables having common oscillation(s) accurately. A new index called the oscillation contribution index (*OCT*) has been proposed to isolate the few key variables as the root cause candidates of plant-wide oscillations.

The steps required to carry out this analysis have been summarized. Two industrial case studies were presented to demonstrate the efficacy of the new procedure.

## Acknowledgements

The authors are grateful for the financial support from the Natural Sciences and Engineering Research Council of Canada (NSERC), Matrikon Inc. and the Alberta Science and Research Authority (ASRA), through the NSERC-Matrikon-ASRA Senior Industrial Research Chair program at the University of Alberta. The authors would like to thank John Cox and Michael Paulonis of Eastman Chemical Company, and Hisato Douke and Souichi Amano of Mitsubishi Chemical Corporation for allowing us the use of their plant data. The authors would also like to thank Dr. Bhushan Gopaluni for pointing out the papers by Stoffer and co-workers for this study.

## References

- Brillinger, D.R. (1981). *Time series: Data analysis and theory (2nd ed., 1981)*. San Francisco: Holden-Day.
- Choudhury, M. A. A. S. (2004). Detection and diagnosis of control loop nonlinearities, valve stiction and data compression. Phd thesis. Department

of Chemical and Materials Engineering, University of Alberta, Canada.

- Choudhury, M. A. A. S., N. F. Thornhill and S. L. Shah (2005a). Modelling valve stiction. *Control Engineering Practice* **13**, 641–658.
- Choudhury, M. A. A. S., S. L. Shah and N. F. Thornhill (2004a). Detection and quantification of control valve stiction. In: *The proceedings of DYCOPS 2004, July 5-7, 2004*. Cambridge, USA.
- Choudhury, M. A. A. S., S. L. Shah and N. F. Thornhill (2004b). Diagnosis of poor control loop performance using higher order statistics. *Automatica* **40**(10), 1719–1728.
- Choudhury, M. A. A. S., S. L. Shah, N. F. Thornhill and David S. Shook (2006). Automatic detection and quantification of stiction in control valves. *Control Engineering Practice, In Press, Available online from 30 November 2005*.
- Choudhury, M. A. A. S., V. Kariwala, S. L. Shah, H. Douke, H. Takada and N. F. Thornhill (2005b). A simple test to confirm control valve stiction. In: *The 16th IFAC World Congress*. Prague, Czech Republic.
- Desborough, L. and R. Miller (2001). Increasing customer value of industrial control performance monitoring - honeywell's experience. In: *Proc. of CPC VI*. Tuscon, Arizona. pp. 172–192.
- Hannan, E.J. (1970). *Multiple time series*. Wiley. New York.
- Hägglund, T. (1995). A control-loop performance monitor. *Control Engng. Prac.* **3**, 1543–1551.
- Jenkins, Gwilym M. and Donald G. Watts (1968). *Spectral Analysis and Its Applications*. Holden Day.
- McDougall, A.J., D.S. Stoffer and D.E. Tyler (1997). Optimal transformations and the spectral envelop for real-valued time series. *Journal of Statistical Planning and Inference* **57**, 195–214.
- Miao, T. and D.E. Seborg (1999). Automatic detection of excessive oscillatory feedback control loops. In: *Proc. of IEEE international Conference on Control Applications*. Kohala Coast, Hawai'i.
- Qin, S.J. (1998). Control performance monitoring - a review and assessment. *Computer and Chemical Engineering* **23**, 173–186.
- Rengaswamy, R., T. Hagglund and V. Venkatasubramanian (2001). A qualitative shape analysis formalism for monitoring control loop performance.. *Engng. Appl. Artificial Intell.* **14**, 23–33.
- Shumway, Robert H. (1988). *Applied Statistical Time Series Analysis*. New Jersey: Prentice Hall.
- Stoffer, David S. (1999). Detecting common signals in multiple time series using the spectral envelope. *Journal of the American Statistical Association* **94**, 1341–1356.

- Stoffer, David S., David E. Tyler and Andrew J. McDougall (1993). Spectral analysis for categorical time series: Scaling and spectral envelope. *Biometrika* **80**, 611–622.
- Stoffer, David S., David E. Tyler and David A. Wendt (2000). The spectral envelope and its applications. *Statistical Science* **15**, 224–253.
- Tangirala, A.K. and S.L. Shah (2005). Non-negative matrix factorization for detection of plant-wide oscillations. Internal report. Department of Chemical and Materials Engineering, University of Alberta, Canada.
- Thornhill, N.F. and T. Hägglund (1997). Detection and diagnosis of oscillation in control loops. *Control Engineering Practice*.
- Thornhill, N.F., B. Huang and H. Zhang (2003a). Detection of multiple oscillations in control loops. *Journal of Process Control* **13**, 91–100.
- Thornhill, N.F., John W. Cox and Michael A. Paulonis (2003b). Diagnosis of plant-wide oscillation through data-driven analysis and process understanding. *Control Engineering Practice* **11**, 1481–1490.
- Thornhill, N.F., S.L. Shah and B. Huang (2001). Detection of distributed oscillations and root-cause diagnosis. In: *Proc. of CHEMFAS 4*. Jeju Island, Korea. pp. 167–172.
- Thornhill, N.F., S.L. Shah, B. Huang and A. Vishnubhotla (2002). Spectral principal component analysis of dynamic process data. *Control Engineering Practice* **10**, 833–846.